

Penerapan *K-Means Clustering* pada Penentuan Jenis Pembelajaran di Universitas Pasir Pengaraian

Luth Fimawahib¹, Erni Rouza²

Universitas Pasir Pengaraian, Jl. Tuanku Tambusai, Rambah Hilir, Rokan Hulu, Riau, Indonesia^{1,2}

E-mail: luthfimawahib@gmail.com¹, ernirouzait@gmail.com²

Abstrack - The Covid-19 pandemic that has hit the world has changed the pattern of human life, including in the process of teaching and learning activities in universities. One of the universities affected by this pandemic is Pasir Pengaraian University. Lectures carried out by Pasir Pengaraian University during the Covid-19 pandemic consist of at least three forms, namely offline, online and blended learning. Efforts to assess which learning method is the most effective become important to measure the level of success of the teaching and learning process, so this study aims to determine the lecture strategy at Pasir Pengaraian University using the K-Means Clustering method. K-Means Clustering algorithm is a method in data mining that can be used for data grouping. CRISP-DM is a data mining methodology used in this study. The research dataset was obtained from the Even semester 2020 lecturer learning reports. RapidMiner was used as a tool to process the data. Clusters were formed as many as 3 (three) with the results of Cluster 1 (49 lecturers), Cluster 2 (17 lecturers), and Cluster 3 (54 lecturers). Based on these results, the lecture strategy with the Blended Learning type of learning is the most appropriate choice to be used at Pasir Pengaraian University, because apart from this Cluster having the highest number of memberships, in this Cluster the highest percentage of places to study are Classrooms/Labors and Meeting Applications, namely blend of offline and online lectures. The blended learning strategy has proven to be representative for use during the pandemic. Evaluation using DBI or Davies-Bouldin Index. The DBI value obtained is -1.163. Cluster evaluation is not good when viewed at this value, because it is negative and not close to zero.

Keywords - *K-Means, Learning, Universitas Pasir Pengaraian*

Intisari – Pandemi Covid-19 yang melanda dunia, merubah pola kehidupan manusia termasuk dalam proses kegiatan belajar mengajar di perguruan tinggi. Salah satu perguruan tinggi yang terdampak pandemi ini adalah Universitas Pasir Pengaraian. Perkuliahan yang dilaksanakan oleh Universitas Pasir Pengaraian dimasa pandemi Covid-19 setidaknya terdiri dari tiga bentuk, yaitu *offline, online* dan *blended learning*. Upaya untuk menilai metode pembelajaran mana yang paling efektif menjadi penting untuk mengukur tingkat keberhasilan proses belajar mengajar, sehingga penelitian ini bertujuan untuk menentukan strategi perkuliahan di Universitas Pasir Pengaraian dengan menggunakan metode *K-Means Clustering*. Algoritma *K-Means Clustering* merupakan suatu metode dalam *data mining* yang dapat dimanfaatkan untuk pengelompokan data. CRISP-DM merupakan metodologi *data mining* yang dipakai dalam penelitian ini. *Dataset* penelitian diperoleh dari laporan pembelajaran dosen semester Genap 2020. RapidMiner digunakan sebagai *tool* untuk memproses data tersebut. *Cluster* yang dibentuk sebanyak 3 (tiga) dengan hasil *Cluster 1* (49 dosen), *Cluster 2* (17 dosen), dan *Cluster 3* (54 dosen). Berdasarkan hasil ini, strategi perkuliahan dengan jenis pembelajaran *Blended Learning* menjadi pilihan yang paling tepat untuk digunakan di Universitas Pasir Pengaraian, karena selain *Cluster 3* ini memiliki jumlah keanggotaan yang terbanyak, pada *Cluster* ini persentase tertinggi tempat belajar adalah Ruang Kelas/Labor dan Aplikasi Meeting, yaitu perpaduan antara kuliah secara *offline* dan *online*. Startegi perkuliahan *blended learning* terbukti representatif digunakan dimasa pandemi. Evaluasi menggunakan DBI atau *Davies-Bouldin Index*. Nilai DBI yang didapatkan sebesar -1.163. Evaluasi *Cluster* kurang baik jika dilihat pada nilai ini, karena bernilai negatif dan tidak mendekati nol.

Kata Kunci - *K-Means, Pembelajaran, Universitas Pasir Pengaraian*

I. PENDAHULUAN

Melonjaknya kasus yang terkonfirmasi positif Covid-19 mengharuskan pemerintah mengeluarkan beberapa surat edaran dan keputusan, seperti Surat Edaran Menteri Pendidikan dan Kebudayaan Nomor 4 Tahun 2020 yang mengatur tentang Pelaksanaan Kebijakan Pendidikan dalam Masa Darurat Penyebaran Covid-19 dan surat keputusan Menteri Pendidikan dan Kebudayaan Nomor 719/P/2020 tentang Pedoman Pelaksanaan Kurikulum Pada Satuan Pendidikan dalam Kondisi Khusus Menteri Pendidikan dan Kebudayaan Republik Indonesia.

Universitas Pasir Pengaraian merupakan perguruan tinggi yang ada di Kabupaten Rokan Hulu Riau. Menindaklanjuti surat edaran dan keputusan menteri yang ada, Universitas Pasir Pengaraian mengeluarkan Surat Edaran Rektor Universitas Pasir Pengaraian Nomor 210/UPP/MN/III/2020 Perihal Pencegahan Penyebaran Corona Virus Disease 2019 (Covid-19) dengan salah satu poinnya adalah menghentikan perkuliahan tatap muka.

Mengganti sistem perkuliahan konvensional dengan perkuliahan secara *online*, memberikan banyak dampak, seperti belum siapnya civitas akademika yang ada, terbatasnya infrastruktur untuk mendukung perkuliahan *online*, dan lain sebagainya. Melihat hal tersebut, pihak perguruan tinggi pada akhirnya menawarkan beberapa solusi antara lain tetap melaksanakan kuliah tatap muka untuk matakuliah praktikum dengan protokol kesehatan yang ketat, memberikan bantuan flashdisk ke mahasiswa yang berisi materi video ajar dari dosen, perkuliahan dengan aplikasi Smart UPP, dan pelaksanaan kuliah dengan bantuan aplikasi *meeting*. Sehingga dari beberapa tipe pembelajaran ini bisa dikelompokkan ke dalam tiga jenis yaitu perkuliahan secara *offline/luring*, *online*, dan *blended learning*.

Sistem pembelajaran di perguruan tinggi harus bisa diukur, sehingga capaian *output* dari proses belajar mengajar bisa diketahui. Oleh karena itu, perguruan tinggi harus mampu menentukan jenis pembelajaran seperti apa yang benar-benar tepat untuk diimplementasikan. Sementara, saat ini Universitas Pasir Pengaraian belum mampu mengklaster jenis pembelajaran sehingga belum bisa mengukur jenis pembelajaran apa yang paling tepat dan relevan untuk diimplementasikan.

Metode penelitian yang digunakan untuk *Clustering* jenis pembelajaran di Universitas Pasir Pengaraian adalah *K-Means*. *K-Mean* merupakan bagian dari metode yang ada dalam *data mining*. Asal data bisa terdiri dari *warehouse*, *database*, *repository*, data dalam sebuah sistem dinamis, dan web dapat ditemukan pola dan pengetahuan barunya dengan *data mining* [1]. Terdapat banyak algoritma dalam *data mining* yang bisa difungsikan untuk media pengelompokan atau *clustering* data, salah satunya adalah *K-Means* ini. Pemilihan metode *K-Mean* sebagaimana tujuan dari penelitian ini, yaitu melakukan pengelompokan strategi jenis pembelajaran di Universitas Pasir Pengaraian, dan *K-Mean* merupakan metode yang bisa digunakan untuk melakukan pengelompokan tersebut. Bila dibandingkan dengan metode *Clustering* lainnya, seperti *Fuzzy C-Means* misalnya, metode *K-Mean* relatif lebih cepat proses iterasinya untuk mendapatkan *cluster*. Perhitungan *K-Mean* tidak serumit metode *Fuzzy C-Means*. Selain itu, hasil *Cluster* pada metode *K-Mean* lebih akurat [2]. Ada beragam cara dalam melakukan *clustering*, salah satu cara tersebut adalah penentuan anggota dalam grup yang sama dengan membuat suatu peraturan berdasar tingkat yang sama diantara bagian-bagiannya.

Penelitian terdahulu yang menggunakan *K-Means Clustering* adalah penelitian yang dilakukan oleh [3] yaitu melakukan *Clustering* perilaku peserta yang mengikuti belajar mengajar daring menggunakan *Log Data Moodle*. Penentuan jumlah *cluster* pada penelitian tersebut menggunakan metode *Elbow* dengan jumlah *cluster* optimal sebanyak 3 *cluster*. Hasil penelitian menunjukkan bahwa aktivitas peserta pelatihan bisa digunakan untuk mengklaster 3 (tiga) kelompok perilaku peserta pelatihan kursus. Penelitian lainnya yang menggunakan metode sama adalah [4] yang mengklaster daerah pengguna teknologi di Indonesia yang dibagi kedalam 3 (tiga) *cluster* yaitu tinggi, sedang dan rendah dengan hasil evaluasi menggunakan *Davies-Bouldin Index* (DBI) yang cukup optimal sebesar 0,222. Jika merujuk pada penelitian

yang dilakukan oleh peneliti sebelumnya dengan metode yang sama, maka pembeda pada penelitian ini adalah pada jenis perilaku objek yang diteliti dan teknologi yang digunakan. Jenis perilaku objek dan teknologi yang digunakan oleh objek yang diteliti lebih bervariasi.

Penelitian ini memiliki tujuan untuk memahami *data mining* dengan metode *K-Means Clustering* dalam mengelompokkan jenis pembelajaran yang ada di Universitas Pasir Pengaraian. Selain itu tujuannya adalah mengidentifikasi jenis pembelajaran yang paling tepat menggunakan metode *K-Means*. Adapun manfaat dari penelitian ini bisa menjadi acuan bagi pihak Universitas Pasir Pengaraian, jenis pembelajaran apa yang paling representatif diimplementasikan, terutama dalam kondisi wabah pandemi. Manfaat bagi peneliti adalah menambah pengetahuan baru tentang konsep klusterisasi menggunakan Algoritma *K-Means Clustering*.

II. SIGNIFIKANSI STUDI

A. Studi Literatur

1. Data Mining

Sebuah pengetahuan dan informasi-informasi penting yang bisa menaikkan keuntungan, menekan biaya pengeluaran, bisa digali dengan proses *data mining* [5]. Beberapa teknik yang digunakan dalam data mining antara lain matematika, statistik, *artificial intelligence*, dan *machine learning* yang fungsinya mengekstrak dan mengidentifikasi informasi dan pengetahuan berguna yang saling terhubung di beragam *database* [6]. Proses dari penambahan data atau *data mining* menghasilkan suatu hal yang bisa dijadikan modal untuk penentuan langkah diwaktu mendatang [7]. Proses menggali ini bisa dianalogikan seperti pengeboran minyak di laut lepas. Jika laut dan segala isi di dalamnya adalah basis data, maka minyak adalah pengetahuan baru.

2. Clustering

Clustering adalah satu dari banyak metode *data mining* yang mempunyai sifat *unsupervised* (tanpa arahan) yaitu karakteristik setiap *Cluster* tidak ditentukan sebelumnya, namun *Cluster* yang terbentuk mengacu pada atribut-atribut yang serupa dari suatu *Cluster*. Oleh sebab itu, dari sekumpulan data dan atribut-atribut yang mirip akan dikelompokkan sesuai karakteristiknya dan direpresentasikan sebagai titik-titik dalam ruang multidimensi [8].

3. Algoritma *K-Means*

K-Means dikelompokkan menggunakan metode analisa kelompok yang diarahkan pada pemartisian. *N* obyek pengamatan kedalam *K* kelompok atau disebut sebagai *Cluster*, dimana setiap obyek pengamatan memiliki sebuah kelompok dengan rata-rata atau *mean* [9]. Algoritma *K-Means* masuk ke dalam penerapan *data mining Clustering* yang dapat digunakan untuk mengelompokkan data menjadi beberapa kelompok. Kelompok-kelompok yang dibentuk ini memiliki kriteria yang telah ditentukan, lalu data-data yang sesuai dengan kelompoknya dikumpulkan menjadi satu *cluster*. Titik pusat atau *Centroid* adalah hal yang dimiliki oleh setiap *cluster* [10]. Adapun tahapan algoritma *K-Means* sebagai berikut [10]:

- a. Memilih berapa (*k*) *Cluster* yang diharapkan dalam *dataset*
- b. Pilih secara acak *Centroid*
- c. Menghitung jarak terpendek tiap data dengan *Centroid*. Gunakan rumus *Euclidean distance*(*d*) untuk menghitung jarak terpendek dengan *Centroid* tersebut. 2 (dua) titik yang tidak sama bisa diukur dengan metode *Euclidean distance* ini[11]. Adapun rumus tersebut adalah:

$$de = \sqrt{(xi - si)^2 + (yi - ti)^2} \quad (1)$$

Keterangan :

(x,y) = koordinat objek

(s, t) = koordinat Centroid

i = banyaknya objek

- d. Menghitung ulang titik *Cluster* dengan keanggotaan *Cluster* terbaru. Rata-rata seluruh data yang ada dalam *cluster* merupakan pusat *cluster*. Rumus untuk menghitung hal tersebut adalah :

$$V_{ij} = \frac{1}{N_i} = \sum_{k=0}^{N_i} X_{kj} \quad (2)$$

Keterangan :

V_{ij} = Centroid rata-rata pada *Cluster* ke - i untuk variabel ke - j

N_i = Jumlah anggota *Cluster* ke - i

i, k = Indeks dari *Cluster*

j = Indeks variabel

X_{kj} = Nilai data ke - k variabel ke - j untuk *Cluster* tersebut

- e. *Cluster* baru (*Centroid* baru) digunakan untuk menghitung ulang setiap objek. Tahap ini merupakan pembukaan awal iterasi baru. Jika *cluster* masih memiliki anggota yang berpindah, maka kembali ke langkah c. Jika anggota *cluster* tidak berpindah *cluster* lagi, proses clustering selesai.

4. *Offline* (luring), *Online* (daring) dan *Blended Learning*

Keperluannya Luring adalah “luar jaringan”, tidak menggunakan jaringan komputer, contohnya tatap muka langsung atau menggunakan buku pegangan siswa. *Online* (daring) merupakan metode pembelajaran yang memanfaatkan metode interaktif berbasis *Learning Manajemen System* (LMS), seperti menggunakan Google Meet, Zoom, Google Drive, dan sebagainya. [12]. Adapun *Blended Learning* adalah gabungan atau kombinasi dari pembelajaran *Offline* dan *Online*.

5. RapidMiner

Ralf Klinkenberg, Ingo Mierswa, dan Simon Fischer merupakan pengembang RapidMiner di *Artificial Intelligence Unit* dari University of Dortmund. Aplikasi ini adalah aplikasi *open source* berlisensi AGPL (*GNU Affero General Public License*) yang dapat digunakan untuk mengolah data mining [8].

6. CRISP-DM

Cross-Industry Standard Process for Data Mining (CRISP-DM) merupakan metode penyelesaian masalah dalam *data mining*. Tahapan CRISP-DM yaitu *business understanding*, *data understanding*, *data preparation*, *modeling*, *evaluation*, dan *deployment* [7].

B. Tempat dan Waktu Penelitian

Penelitian berada di Universitas Pasir Pengaraian. Waktu penelitian yaitu Mei 2021 sampai September 2021.

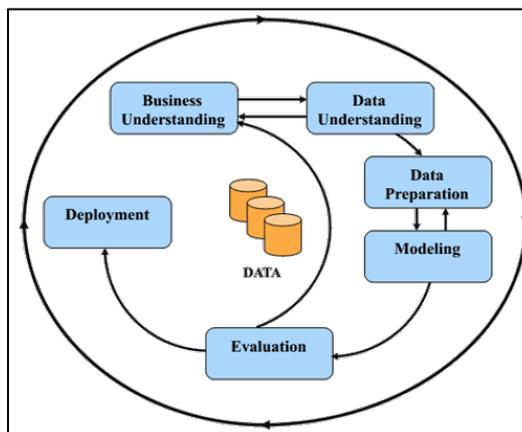
C. Data

Penelitian ini menggunakan data hasil pelaksanaan pengajaran dosen Universitas Pasir Pengaraian pada semester genap 2020. Data ini berjumlah 120, terdiri dari beberapa variabel

seperti no, nama dosen, fakultas, prodi, metode pembelajaran di ruang kelas/labor, kelas/labor dan aplikasi *meeting*, dan aplikasi *meeting*.

D. Metode dan Tahap Penelitian

Algoritma *K-Means Clustering* adalah metode yang diimplementasikan untuk pengelompokan metode pembelajaran di Universitas Pasir Pengaraian. Pengujian atau implementasi metode menggunakan RapidMiner. *Cluster* yang akan dibentuk sebanyak 3 (tiga), yaitu *offline/luring*, *blended learning* dan *online* sebagaimana uraian pada pendahuluan. Tahapan penelitian menggunakan metodologi CRISP-DM dengan penjelasan langkah-langkahnya sebagai berikut:



Gambar 1. Alir Siklus CRISP-DM (<ftp://public.dhe.ibm.com>)

1. Tahap *Business Understanding* : Memahami objek bisnis yang akan dikembangkan informasi dan pengetahuan barunya dengan *data mining*. Penelitian ini menerapkan metode *K-Means Clustering* untuk pengelompokan jenis pembelajaran.
2. Tahap *Data Understanding* : Setelah memahami objek bisnis dengan baik, tahap ini adalah pemahaman terhadap data. Dimulai dari data apa saja yang dikumpulkan, mendesripsikan data, eksploarsi data serta memeriksa dengan cermat kualitas data yang akan diolah.
3. Tahap *Data Preparation* : Setelah tahap dua dilaksanakan, maka pada tahap ini melakukan identifikasi dan penyusunan jawaban terhadap data yang telah diperoleh agar bisa mengelompokkan dan memilah data ke dalam kelompok-kelompok yang sudah ditentukan.
4. Tahap *Modeling* : Tahap ini melakukan penentuan model yang hendak dipakai dalam pengelompokan jenis pembelajaran. *K-Means Clustering* merupakan model yang dipilih pada penelitian ini.
5. Tahap *Evaluation* : Tahap ini merupakan proses pengujian terhadap kelompok yang dibentuk menggunakan algoritma *K-Means Clustering* dengan mengelompokkan secara manual.
6. Tahap *Deployment* : Laporan hasil pengolahan data yang telah selesai, yaitu berupa pengetahuan yang diperoleh

III. HASIL DAN PEMBAHASAN

A. *Business Understanding (Pemahaman Bisnis)*

Universitas Pasir Pengaraian (UPP) sebagai perguruan tinggi yang cukup berkembang pesat untuk ukuran di kabupaten, harus mampu menghadapi segala bentuk perubahan, baik dari

sisi perubahan teknologi maupun perubahan kebijakan pemerintah. Tujuan dari penelitian ini adalah untuk melihat jenis pembelajaran yang paling representatif digunakan oleh civitas UPP. Kegiatan observasi dilakukan secara langsung di UPP. *Output* dari penelitian ini diharapkan memiliki nilai guna bagi UPP dalam menentukan jenis pembelajaran.

B. Data Understanding (Pemahaman Data)

Ketika penelitian sudah ditentukan tujuannya, maka selanjutnya adalah memahami data yang hendak diolah. Penelitian ini menggunakan *dataset* awal berupa laporan pengajaran dosen UPP semester genap 2020. Data ini dikumpulkan, dideskripsikan datanya, dilakukan eksplorasi data dan diverifikasi kualitas datanya.

TABEL I
DESKRIPSI DATA PENGAJARAN DOSEN

Atribut	Tipe Data	Keterangan
No	Numerik	Nomor urut nama dosen
Nama Dosen	Nominal	Dosen di Universitas Pasir Pengaraian
Fakultas	Nominal	Fakultas di Universitas Pasir Pengaraian
Prodi	Nominal	Prodi di Universitas Pasir Pengaraian
Ruang Kelas/Labor	Numerik	Persentase menggunakan Kelas/Labor
Ruang Kelas/Labor dan Aplikasi Meeting	Numerik	Persentase menggunakan Kelas/Labor dan Aplikasi <i>meeting</i>
Aplikasi Meeting	Numerik	Persentase menggunakan aplikasi <i>meeting</i>

Tabel I mendeskripsikan data pengajaran dosen UPP semester genap 2020. Dari Tabel I ini bisa dipahami atribut, tipe data, dan keterangan. Setelah data dideskripsikan, agar mempermudah memahami data supaya diperoleh pengetahuan atau pemahaman data, maka langkah berikutnya adalah melakukan proses *explore the data*. Dari data yang diexplore ini, langkah selanjutnya adalah memverifikasi kualitas data dengan cara mengecek data secara manual dengan *Microsoft Excel* menggunakan fitur *Conditional Formatting*. Hal ini masih bisa dilakukan, mengingat jumlah data yang masih terjangkau untuk dilakukan verifikasi. Jika tidak terdapat nilai yang hilang dan duplikasi data, maka kualitas data adalah baik.

TABEL II
EXPLORE DATA PENGAJARAN DOSEN

Atribut	Tipe Data	Keterangan
No	Numerik	Nomor urut nama dosen Genap 2020 : 1-120
Nama Dosen	Nominal	Nama Dosen di Universitas Pasir Pengaraian Genap 2020 : Terdapat 120 dosen
Fakultas	Nominal	Nama Fakultas di Universitas Pasir Pengaraian Genap 2020 : Terdapat 7 Fakultas
Prodi	Nominal	Nama Prodi di Universitas Pasir Pengaraian Genap 2020 : Terdapat 18 Prodi
Ruang Kelas/Labor	Numerik	Persentase menggunakan Kelas/Labor Genap 2020 : 120 persentase
Ruang Kelas/Labor dan Aplikasi Meeting	Numerik	Persentase menggunakan Kelas/Labor dan Aplikasi <i>meeting</i> Genap 2020 : 120 persentase
Aplikasi Meeting	Numerik	Persentase menggunakan aplikasi <i>meeting</i> Genap 2020 : 120 persentase

Pada Tabel II Explore Data Pengajaran Dosen di atas tidak ditemukan *missing value* dan duplikasi data.

C. Data Preparation (Persiapan Data)

Segala aktivitas dalam mengolah data awal dengan tujuan menciptakan *dataset final* adalah konsep dari *Data Preparation*. Atribut data yang tidak diperlukan pada saat pemodelan

dibersihkan terlebih dahulu. Proses pembersihan dengan cara menghapus atau menghilangkan kolom-kolom yang tidak dibutuhkan saat pemodelan. Kolom yang dihilangkan tersebut adalah kolom Fakultas dan Prodi. Tabel III berikut ini merupakan data pengajaran dosen genap 2020 setelah proses *cleaning* yang menyisakan 4 atribut.

TABEL III
EXPLORE DATA PENGAJARAN DOSEN CLEANING

No	Nama Dosen	Ruang Kelas/Labor	Ruang Kelas/Labor dan Aplikasi Meeting	Aplikasi Meeting
1	RIVI ANTONI, S.Pd, M.Pd	54,35	76,45	60,76
2	ERIPUDDIN, S.Hum., M.Pd	67,31	87,15	56,8
3	PIPIT RAHAYU, M.Pd	65,93	88,45	60,88
4	ANDRI DONAL, S.Pd, M.Pd	70,56	67,76	40,56
5	EVI KASYULITA, S.Pd., M.Pd	83,12	70,16	60,34
6	UMMI RASYIDAH, M.Pd	75,45	67,66	45,86
7	BATDAL NIATI, M.Pd	49,56	67,89	43,16
8	AZMI ASRA, S.Si, M.Pd	67,88	68,67	43,56
9	IKA DARUWATI, M.Sc	76,32	75,45	56,17
10	YEZA FEBRIANI, M.Sc	50,12	76,45	56,76
11	RINDI GENESA HATIKA, M.Sc	52,45	66,51	57,86
12	SOHIBUN, M.Pd	83,55	56,72	45,46
13	Dr. DEDI MARDIANSYAH, M.Si	65,42	57,82	56,67
14	NURHIKMAH SASNAJUNAI, M.Si	50,14	67,84	40,34
15	MUARIF ARHAS PUTRA, M.Pd	59,47	78,93	60,45
16	LOLIA MANURIZAL, M.Pd	51,66	65,76	61,34
17	ARDO YULPIKO PUTRA, S.Pd., M.Kes	32,45	15,98	78,79
18	MUHAMMAD NURHISYAM ALI SETIAWAN, M.Or	12,54	34,56	80,17
19	RIDWAN SINURAT, M. Or	89,43	71,45	50,16
...
120	ERMIZA, SST. M BIOMED	84,71	80,67	40,56

Penerapan *Algoritma K-Means*

Tahapan yang dilakukan dalam mengelompokkan data menjadi beberapa *Cluster* dengan menggunakan *Algoritma K-Means* antara lain :

- a. Memilih berapa jumlah *Cluster*. *Cluster* dalam penelitian ini berjumlah 3 (tiga).
- b. Memilih secara acak pusat *Cluster* (centroid) awal. Dari Tabel III, dipilih secara acak centroid awal sebagaimana terlihat pada Tabel IV:

TABEL IV
CENTROID AWAL DATA DOSEN

C1	70,56	67,76	40,56
C2	50,16	78,93	74,56
C3	30,16	78,65	85,66

- c. Menghitung jarak terpendek tiap data dengan *Centroid*. Gunakan rumus *Euclidean distance* (d) untuk menghitung jarak terpendek dengan *Centroid* tersebut:

$$C1 = \sqrt{(54,35 - 70,56)^2 + (76,45 - 67,76)^2 + (60,76 - 40,56)^2}$$

$$C1 = 27,32$$

$$C2 = \sqrt{(54,35 - 50,16)^2 + (76,45 - 78,93)^2 + (60,76 - 74,56)^2}$$

$$C2 = 14,63$$

$$C3 = \sqrt{(54,35 - 30,16)^2 + (76,45 - 78,65)^2 + (60,76 - 85,66)^2}$$

$$C3 = 34,79$$

Perhitungan dilakukan sampai dengan data dosen ke 120. Tabel V menunjukkan hasil perhitungan jarak *Cluster* untuk iterasi 1.

TABEL V
HASIL ITERASI 1 DATA DOSEN

No	Nama Dosen	C1	C2	C3	Cluster
1	RIVI ANTONI, S.Pd, M.Pd	27,32	14,63	34,79	2
2	ERIPUDDIN, S.Hum., M.Pd	25,50	26,02	47,80	1
3	PIPIT RAHAYU, M.Pd	29,37	22,94	44,60	2
4	ANDRI DONAL, S.Pd, M.Pd	0,00	41,19	61,52	1
5	EVI KASYULITA, S.Pd., M.Pd	23,55	36,95	59,31	1
6	UMMI RASYIDAH, M.Pd	7,21	39,88	61,29	1
7	BATDAL NIATI, M.Pd	21,16	33,29	47,94	1
8	AZMI ASRA, S.Si, M.Pd	4,12	37,15	57,40	1
9	IKA DARUWATI, M.Sc	18,33	32,17	54,87	1
10	YEZA FEBRIANI, M.Sc	27,49	17,97	35,19	2
11	RINDI GENESA HATIKA, M.Sc	25,08	20,94	37,64	2
12	SOHIBUN, M.Pd	17,74	49,55	70,34	1
13	Dr. DEDI MARDIANSYAH, M.Si	19,62	31,60	50,18	1
14	NURHIKMAH SASNAJUNAI, M.Si	20,42	35,97	50,69	1
15	MUARIF ARHAS PUTRA, M.Pd	25,36	16,90	38,66	2
16	LOLIA MANURIZAL, M.Pd	28,16	18,72	34,93	2
17	ARDO YULPIKO PUTRA, S.Pd., M.Kes	74,80	65,53	63,09	3
18	MUHAMMAD NURHISYAM ALI SETIAWAN, M.Or	77,70	58,44	47,80	3
19	RIDWAN SINURAT, M. Or	21,49	46,83	69,46	1
...
120	ERMIZA, SST. M BIOMED	19,15	48,50	70,81	1

- d. Menghitung ulang titik *Cluster* dengan keanggotaan *Cluster* terbaru. Rata-rata seluruh data yang ada dalam *Cluster* merupakan pusat *cluster*.

TABEL VI
CENTROID BARU

C1	71,77	64,94	43,70
C2	56,60	74,06	66,54
C3	34,49	57,06	76,26

Perhitungan dilakukan sampai dengan data dosen ke 120. Tabel VII menunjukkan hasil perhitungan jarak *Cluster* untuk iterasi 2.

TABEL VII
HASIL ITERASI 2 DATA DOSEN

No	Nama Dosen	C1	C2	C3	Cluster
1	RIVI ANTONI, S.Pd, M.Pd	25,09	8,71	37,35	2
2	ERIPUDDIN, S.Hum., M.Pd	24,70	22,92	54,00	2
3	PIPIT RAHAYU, M.Pd	28,15	21,32	52,11	2
4	ANDRI DONAL, S.Pd, M.Pd	4,79	32,44	58,48	1

No	Nama Dosen	C1	C2	C3	Cluster
5	EVI KASYULITA, S.Pd., M.Pd	20,36	29,89	59,07	1
6	UMMI RASYIDAH, M.Pd	5,31	31,13	58,76	1
7	BATDAL NIATI, M.Pd	21,21	25,99	43,84	1
8	AZMI ASRA, S.Si, M.Pd	4,21	28,43	54,72	1
9	IKA DARUWATI, M.Sc	15,93	25,30	56,14	1
10	YEZA FEBRIANI, M.Sc	25,97	13,07	37,05	2
11	RINDI GENESA HATIKA, M.Sc	22,31	12,06	33,78	2
12	SOHIBUN, M.Pd	15,80	40,15	64,71	1
13	Dr. DEDI MARDIANSYAH, M.Si	15,10	21,52	43,40	1
14	NURHIKMAH SASNAJUNAI, M.Si	21,06	28,64	46,51	1
15	MUARIF ARHAS PUTRA, M.Pd	23,22	11,58	42,36	2
16	LOLIA MANURIZAL, M.Pd	25,07	9,86	30,77	2
17	ARDO YULPIKO PUTRA, S.Pd., M.Kes	71,27	61,20	41,36	3
18	MUHAMMAD NURHISYAM ALI SETIAWAN, M.Or	74,74	57,48	28,04	3
19	RIDWAN SINURAT, M. Or	20,37	39,48	68,98	1
...
120	ERMIZA, SST. M BIOMED	21,00	42,00	72,34	1

Perhitungan yang dilakukan terhadap nilai centroid dilakukan secara terus-menerus sampai proses iterasi *Cluster* sudah tidak berubah lagi nilainya. Pada iterasi ke tujuh di dalam penelitian ini, proses berhenti. Tabel VIII menampilkan hasil iterasi terakhir yaitu iterasi ke tujuh dengan jumlah *Cluster* pertama sebanyak 49 *items*, *Cluster* kedua sebanyak 17 *items*, dan *Cluster* ketiga sebanyak 54 *items*.

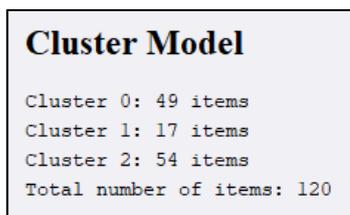
TABEL VIII
HASIL ITERASI 7 DATA DOSEN

No	Nama Dosen	C1	C2	C3	Cluster
1	RIVI ANTONI, S.Pd, M.Pd	29,27	3,92	29,39	2
2	ERIPUDDIN, S.Hum., M.Pd	27,83	17,36	45,77	2
3	PIPIT RAHAYU, M.Pd	31,40	16,59	44,03	2
4	ANDRI DONAL, S.Pd, M.Pd	5,75	27,10	50,04	1
5	EVI KASYULITA, S.Pd., M.Pd	20,53	26,63	49,45	1
6	UMMI RASYIDAH, M.Pd	4,92	26,13	49,77	1
7	BATDAL NIATI, M.Pd	24,75	22,24	37,49	2
8	AZMI ASRA, S.Si, M.Pd	7,78	23,06	46,28	1
9	IKA DARUWATI, M.Sc	17,64	20,50	46,84	1
10	YEZA FEBRIANI, M.Sc	30,25	9,47	29,95	2
11	RINDI GENESA HATIKA, M.Sc	26,26	10,63	25,62	2
12	SOHIBUN, M.Pd	12,19	36,51	55,60	1
13	Dr. DEDI MARDIANSYAH, M.Si	17,18	19,83	34,33	1
14	NURHIKMAH SASNAJUNAI, M.Si	24,33	24,63	40,25	1
15	MUARIF ARHAS PUTRA, M.Pd	27,19	5,58	34,06	2
16	LOLIA MANURIZAL, M.Pd	28,92	10,49	22,30	2
17	ARDO YULPIKO PUTRA, S.Pd., M.Kes	72,78	65,51	43,02	3
18	MUHAMMAD NURHISYAM ALI SETIAWAN, M.Or	77,54	62,27	35,30	3
19	RIDWAN SINURAT, M. Or	18,61	35,02	59,49	1
...
120	ERMIZA, SST. M BIOMED	21,00	42,00	72,34	1

D. Modeling (Pemodelan)

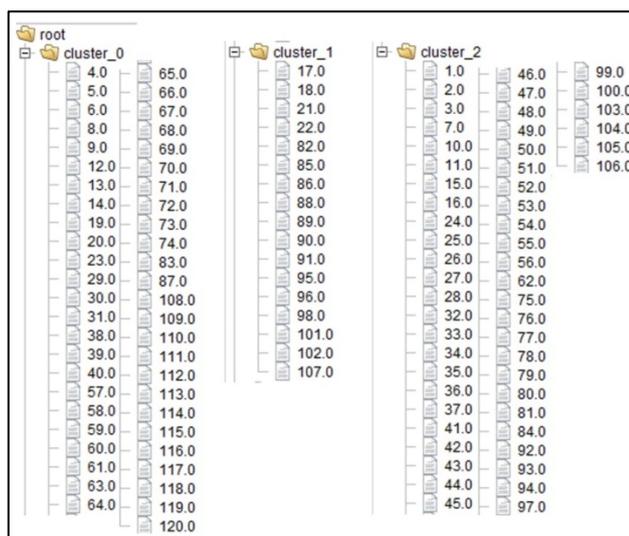
Setelah persiapan data dan pengujian secara manual selesai, tahap berikutnya adalah *modeling* (pemodelan). *Tool* yang digunakan yaitu RapidMiner menggunakan metode

algoritma *K-Means Clustering*. Dari 120 data pengajaran dosen, dibentuklah 3 (tiga) *Cluster* yaitu sistem pengajaran secara *offline*, *online* dan *blended learning*.



Gambar 1. Cluster Model dalam RapidMiner

Gambar 1 merupakan tampilan hasil *cluster* (kelompok) ketika data diuji dengan RapidMiner. *Cluster* pertama sebanyak 49 *items*, *Cluster* kedua sebanyak 17 *items*, dan *Cluster* ketiga sebanyak 54 *items*.



Gambar 2. Tampilan Cluster dalam Folder di RapidMiner

Dapat disimpulkan bahwa *Cluster* jenis pembelajaran dosen di Universitas Pasir Pengaraian adalah :

- Cluster 1 (*Offline*) adalah Andri Donal, S.Pd, M.Pd, Evi Kasyulita, S.Pd,. M.Pd, Umni Rasyidah, M.Pd, Azmi Asra, S.SI, M.Pd, Ika Daruwati, M.Sc, Sohibun, M.Pd, Dr. Dedi Mardiansyah, M.SI, Nurhikmah Sasnajunaidi, M.SI, Ridwan Sinurat M. OR, Masji Janiarli, M.Kes, Ria Karno, S.Pd,.M.SI, Rena Lestari, M.Pd, Dahlia, S.Pd, Lusi Eka Afri, M.SI, Welfen Aida, S.Pd,. M.Pd, Cicilia Melinda, M.Pd, Rina Ari Rohmah, M.Pd, Afrijal, SE., MM, Nofrianti, SE,.M.SI, Zulkarnain, M . Ak, Bambang Edison S.Pd, MT, Alfi Rahmi, ST. M.Eng, Rismalinda, ST,.MT, DR. Pada Lumba, ST, MT, Khairul Fahmi, MT, Juli Marliansyah, MT, Arie Syahrudin S, MT, Arifal Hidayat, ST. MT, Saiful Anwar, MT, Ahmad Fathoni, MT, Yose Rizal, ST, MT, Heri Suropto, MT, Aprizal, MT, Purwo Subekti, ST, MT, Laily Fitriana,SP,.M.SI, Dr. Defi Delwina, SP,.M.SC, Eka Yuli Handayani, SST,. M.Kes, Andria, SST,.M.Biomed, Anastasia Marliza, SST, Sri Wulandari, SKM,.M.PH, Rahmi Fitria, S.ST,.M.Biomed, Heny Sepduwiana, SST,. M.Kes, Nana Aldriana, SST,.M.Kes, Rika Herawati, SST,.M.Kes, Elvira Junita, SST,. M.Kes, Andriana, SST, NS.Romy Wahyuny, S.Kep,.M.Kes, Yuyun Bewelli Fahmi, SST, Ermiza, SST. M Biomed.

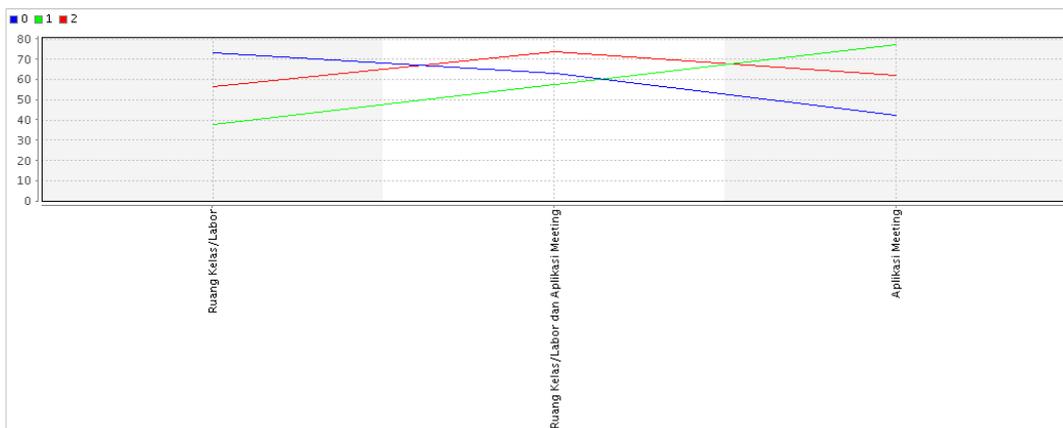
2. Cluster 2 (*Online*) adalah Ardo Yulpiko Putra, S.Pd., M.Kes, Muhammad Nurhisyam Ali Setiawan, M.Or, Aluwis, M.Pd, Riwaldi Putra, M.Pd, Khusnu Abdullah Siregar, M.P, M. Yunan Nasution, MMA, Ikhsan Gunawan, SP, M.MA, Sischa Febriani Yamesa Away, M.App.Sc, Abdul Latif, MH, Zulkifli, SH.MH, Rise Karmilia, SH., M.Hum, Kiki Yasdomi, S.Kom., M.Kom, Hendri Maradona, S.KOM, M.Kom, Dr. Hendri Kurniawan, SE., MM, Basorudin, S.Pd., M.Kom, Budi Yanto, M.Kom, Imam Rangga Bakti, M.Kom.
3. Cluster 3 (*Blended Learning*) adalah Rivi Antoni, S.Pd, M.Pd, Eripuddin, S.Hum., M.Pd, Pipit Rahayu, M.Pd, Batdal Niati, M.Pd, Yeza Febriani, M.Sc, Rindi Genesa Hatika, M.Sc, Muarif Arhas Putra, M.Pd, Lolia Manurizal, M.Pd, Arief Anthonius Purnama, M.Si, Eti Meirina Brahmana, S.Si., M.Si, Jismi Mubarrak, S.Si., M.Si, Filza Yulina Ade, M.Si, Nurul Afifah, S.Pd., M.Pd, Hera Deswita, M.Pd, Arcat, M.Pd, Ratri Isharyadi, M.Pd, Annajmi, M.Pd, Marfi Ario, M.Pd, Nurrahmawati, S.Pd., M.Pd, Dr. Hardianto, S.Pd, M.Pd, Romika Rahayu, M. Pd, Ilham Rahmawati, M. Pd, Ryan Prayogi, M. Pd, Yulfita 'Aini, SE, MM, Seprini, MM, Heffi Christya Rahayu, M.Si, Arrafiqur Rahman, SE, MM, Hidayat, SE., MM, Makmur, SE., M.MA, Hamdi Sari Maryoni, SE.M.EC. DEV, Andi Afrizal, SE., M., Si, AK, CA, Purwantoro, SE, M.Si, Sri Yunawati, ST., M.Acc, Nurhayati, SE, M.Si, Arma Yuliza, SE, M.Si, Anton Ariyanto, ST., M.Eng, Rizah Rizwana Wahyuni, STP.M.Sc, Zulputra, SP., M.Si, Lufita Nur Alfiah, SP., M.Si, Dr. Ir. Edward Bahar, MP, Yuliana Susanti, M.Si, Al Muzafri, M. Si, Fajar Syukron, M. Si, Rina Febrinova, SE., MMA, Nofrizal, Lc., MH, Almadison, MH, Hendri, MH, Mi'rajul Rifqi, M.Cs, Khairul Sabri, M.Kom, Urfi Utami, M.Kom, Erni Rouza, M. Kom, Luth Fimawahib, M. Kom, Satria Riki Mustafa, S.Pd., M.Si, Asep Supriyanto, M.Kom

Hasil perhitungan dengan angka paling kecil ataupun mendekati nilai titik pusat (*centroid*) adalah dasar penentuan sebuah *cluster*.

TABEL IX
HASIL CENTROID AKHIR

Atribut	Cluster 1	Cluster 2	Cluster 3
Ruang Kelas/Labor	73.987	38.273	57.005
Ruang Kelas/Labor dan Aplikasi Meeting	63.882	58.588	74.628
Aplikasi Meeting	43.062	77.674	63.000

Centroid final pada penelitian ini sebagaimana Tabel IX. Berdasarkan Tabel ini, dapat dilihat bahwa *Cluster 1* dengan jumlah dosen sebanyak 49 orang persentase tertinggi tempat belajar adalah pada atribut Ruang Kelas/Labor sebanyak 73,987, disusul dengan Ruang Kelas/Labor dan Aplikasi Meeting sebesar 63,882, kemudian Aplikasi Meeting sebesar 43,062. Hal ini menunjukkan bahwa keaktifan jenis pembelajaran pada *Cluster 1* ada pada Ruang Kelas/Labor. *Cluster 2* dengan jumlah dosen sebanyak 17 orang, persentase tertinggi tempat belajar adalah pada atribut Aplikasi Meeting sebanyak 77,674, disusul dengan Ruang Kelas/Labor dan Aplikasi Meeting sebesar 58,588, lalu Ruang Kelas/Labor sebesar 38,273. Hal ini menunjukkan bahwa jenis pembelajaran pada *Cluster 2* condong ke Aplikasi Meeting. Sedangkan pada *Cluster 3* dengan jumlah dosen sebanyak 54 orang, persentase tertinggi tempat belajar adalah pada atribut Ruang Kelas/Labor dan Aplikasi Meeting sebanyak 74.628, disusul dengan Aplikasi Meeting sebesar 63,000 lalu Ruang Kelas/Labor sebesar 57,005. Hal ini menunjukkan bahwa *Cluster 3* lebih condong pada jenis pembelajaran Ruang Kelas/Labor dan Aplikasi Meeting.



Gambar 3. Grafik hasil K-Means Clustering

Gambar 3 merupakan grafik hasil K-Means Clustering pada *Centroid final*. Garis warna biru merupakan representasi Cluster 1, garis warna hijau representasi Cluster 2 dan sedangkan garis warna merah representasi Cluster 3. Cluster 1 garis warna biru, puncak tertinggi grafik ada pada atribut Ruang Kelas/Labor, Cluster 2 garis warna hijau, puncak tertinggi grafik ada pada atribut Aplikasi Meeting, sedangkan Cluster 3 garis warna merah, puncak tertinggi grafik ada pada atribut Ruang Kelas/Labor dan Aplikasi Meeting.

E. Evaluation (Evaluasi)

```

PerformanceVector

PerformanceVector:
Avg. within centroid distance: -302.431
Avg. within centroid distance_cluster_0: -267.823
Avg. within centroid distance_cluster_1: -456.893
Avg. within centroid distance_cluster_2: -285.206
Davies Bouldin: -1.163
    
```

Gambar 4. Hasil uji performance

Evaluasi pada tahap ini menggunakan DBI atau *Davies-Bouldin Index*. Kinerja pengelompokan Cluster pada umumnya dievaluasi dengan DBI. Adapun hasil uji performance operator pada aplikasi RapidMiner menggunakan *Cluster distance performance*. Seberapa baik kinerja dari *centroid* yang dihasilkan, juga bisa diukur dengan operator ini. Dari hasil implementasi menggunakan RapidMiner untuk menguji data jenis pembelajaran, nilai DBI yang didapatkan sebesar -1,163 sebagaimana Gambar 4 di atas. Evaluasi Cluster kurang baik jika dilihat pada nilai ini, karena bernilai negatif dan tidak mendekati nol. Semakin kecil DBI dan tidak bernilai negatif, semakin baik evaluasi sebuah Cluster.

F. Deployment (Penyebaran)

Pada tahap ini adalah melakukan presentasi atau demo aplikasi yang digunakan, menggunakan aplikasi RapidMiner. Pengujian data berbasis aplikasi desktop. Tahap ini

melibatkan user dalam menggunakan aplikasi data mining dengan metode *K-Means Clustering* untuk penentuan strategi perkuliahan di Universitas Pasir Pengaraian.

IV. KESIMPULAN

Penelitian ini adalah menentukan strategi perkuliahan yang ada di Universitas Pasir Pengaraian dengan menerapkan algoritma *K-Means Clustering*. Jumlah *k* yang digunakan sebanyak 3 (tiga). Dari hasil implementasi pengujian menggunakan RapidMiner terhadap 120 data dosen yang melakukan pembelajaran pada semester genap 2020 diperoleh *Cluster 0* (49 dosen), *Cluster 1* (17 dosen), dan *Cluster 2* (54 dosen). Evaluasi menggunakan DBI atau *Davies-Bouldin Index* dan diperoleh hasil -1,163. Evaluasi *Cluster* kurang baik jika dilihat pada nilai ini, karena bernilai negatif dan tidak mendekati nol. Semakin kecil DBI dan tidak bernilai negatif, semakin baik evaluasi sebuah *Cluster*. Jenis pembelajaran yang mempunyai jumlah tertinggi adalah *Cluster 2* (54 dosen) yaitu pembelajaran dengan metode *Blended Learning*. Strategi perkuliahan dengan jenis pembelajaran *Blended Learning* menjadi pilihan yang paling tepat untuk digunakan di Universitas Pasir Pengaraian, karena selain *Cluster* ini memiliki jumlah keanggotaan yang terbanyak, pada *Cluster* ini persentase tertinggi tempat belajar adalah Ruang Kelas/Labor dan Aplikasi Meeting, yaitu perpaduan antara kuliah secara *offline* dan *online*.

REFERENSI

- [1] E. Rouza and L. Fimawahib, "Implementasi Fuzzy C-Means Clustering dalam Pengelompokan UKM Di Kabupaten Rokan Hulu," *Techno.COM*, vol. 19, no. 4, pp. 481–495, 2020.
- [2] Butarbutar and Dkk, "Komparasi Kinerja Algoritma Fuzzy C-Means dan K-Means dalam Pengelompokan Data Siswa Berdasarkan Prestasi Nilai Akademik Siswa (Studi Kasus : SMP Negeri 2 Pematangsiantar)," *J. Ris. Inf. Tek. Inform.*, vol. 1, no. 2012, pp. 46–55, 2016.
- [3] E. Ikhsan, "Penerapan K-Means Clustering dari Log Data Moodle untuk Menentukan Perilaku Peserta pada Pembelajaran Daring," *Sistemasi*, vol. 10, no. 2, p. 414, 2021, doi: 10.32520/stmsi.v10i2.1285.
- [4] S. F. Mandang and B. N. Sari, "Penerapan K-Means Cluster Pada Daerah Penggunaan Teknologi di Indonesia," *JOINS (Journal Inf. Syst.*, vol. 6, no. 1, pp. 131–138, 2021, doi: 10.33633/joins.v6i1.4545.
- [5] H. Rofiq, K. C. Pelangi, and Y. Lasena, "Penerapan Data Mining Untuk Menentukan Potensi Hujan Harian Dengan Menggunakan Algoritma Naive Bayes," *J. Manaj. Inform. dan Sist. Inf.*, vol. 3, no. 1, pp. 8–15, 2020, [Online]. Available: <http://mahasiswa.dinus.ac.id/docs/skripsi/jurnal/19417.pdf>.
- [6] S. Darma, S. Defit, D. Hartama, and W. Robiansyah, "Penerapan Metode K-Means Dalam Pengelompokan Jumlah Wisatawan Asing Di Indonesia," vol. 2, pp. 255–261, 2020.
- [7] Y. P. Sari, A. Primajaya, and A. S. Y. Irawan, "Implementasi Algoritma K-Means untuk Clustering Penyebaran Tuberkulosis di Kabupaten Karawang," *INOVTEK Polbeng - Seri Inform.*, vol. 5, no. 2, p. 229, 2020, doi: 10.35314/isi.v5i2.1457.
- [8] R. S. and N. Tes., "Penerapan Data Mining Menggunakan Algoritma K-Means Clustering

- Untuk Menentukan Strategi Promosi Mahasiswa Baru (Studi Kasus : Politeknik Lp3i Jakarta),” *J. Lentera Ict*, vol. 3, no. 1, pp. 76–92, 2016.
- [9] D. Meilvinasvita, Safaruddin, and Yuliana, “Vocational education and technology journal,” *Vocat. Educ. Technol. J.*, vol. 1, no. 2, pp. 21–27, 2020, [Online]. Available: <http://ojs.aknacehbarat.ac.id/index.php/vocatech/index>.
- [10] N. Dwitri, J. A. Tampubolon, S. Prayoga, F. Ilmi Zer, and D. Hartama, “Penerapan Algoritma K-Means Dalam Menentukan Tingkat Kepuasan Pembelajaran Online Pada Masa Pandemi Covid-19 di Indonesia,” *Jti (Jurnal Teknol. Informasi)*, vol. 4, no. 1, pp. 101–105, 2020.
- [11] A. Aditya, I. Jovian, and B. N. Sari, “Implementasi K-Means Clustering Ujian Nasional Sekolah Menengah Pertama di Indonesia Tahun 2018/2019,” *J. Media Inform. Budidarma*, vol. 4, no. 1, p. 51, 2020, doi: 10.30865/mib.v4i1.1784.
- [12] A. Malyana, “Pelaksanaan Pembelajaran Daring Dan Luring Dengan Metode Bimbingan Berkelanjutan Pada Guru Sekolah Dasar Di Teluk Betung Utara Bandar Lampung,” *Pedagog. J. Ilm. Pendidik. Dasar Indones.*, vol. 2, no. 1, pp. 67–76, 2020, doi: 10.52217/pedagogia.v2i1.640.

UCAPAN TERIMA KASIH

Ucapan terima kasih peneliti sampaikan kepada Universitas Pasir Pengaraian dan kepada Tim *Jurnal Informatika Polbeng* yang telah menyisihkan waktu demi perbaikan jurnal ini.